Course Assignment: WID3006

Overall mark allocation: 20%

Contents

1   Preliminary Notes: Lovathon

The theme of the Machine Learning Hackathon is **Love**.

Choose, *define*, and *design* a specific love problem that can be solved using Machine Learning.

Examples:

- How compatible are you and your partner based on facial features similarity?
  - Compatibility test based on facial features.
- Can you classify whether a pair of individuals would fall in love based on their personal backgrounds?
  - Love classifier based on features from pairs of couples.
- How much are you in love with your partner? Is it possible to quantify and predict the love for someone?
  - Love value predictor (regression)
- Etc.

Marks shall be awarded based on a specific rubric (Section 2 of this document), aimed at evaluating the overall level of understanding of the students behind the machine learning techniques used to solve the problems.

Machine learning techniques are not limited to the techniques presented in class (i.e. can use other methods such as XGBoost or Random Forest).

IMPORTANT:

Note that, there are **limited** datasets for the different problems surrounding love. And so, you may need to develop or collect your own datasets for the problems that you have chosen. There is even a possibility of developing a **pseudo-dataset** to demonstrate your proposed problem & and solution.

2   Rubric

Choice of assignment

Criteria/ Marks Allocations	1-8 marks	9-15 marks	16-25 marks
Validity of the problem from a machine learning aspect, and the choice of data to support the problem.	Minimally valid with limited data (5)	Adequately valid and convincing data(15)	Highly valid and convincing choice/design of data (25)
Initial flowchart of the proposed solution, gantt chart and tasks distribution among group members for the project (25 marks)	Showcases bare minimal understanding of the problem.	Showcases clear understanding of the problem	Showcases very clear understanding of the problem and well-designed flowchart
Report Evaluation 1:			

(Structure and Clarity, 25 marks) • **Introduction to Problem** • **Hypothesis made for the problem** • **Project Objectives** • **Methodology** • **Elaboration on Data & Features Used** • **Results and discussions** • **Suggestion for future works**	Report meets the bare minimal standard of structure and clarity. Contains minimal data exploration Methodology lacks clarity and requires refinement. Hypothesis is validated but not in entirety via the results and discussions.	Report is well written in terms of format and presentation. Contains enough data exploration. (e.g. plot of histogram, correlation matrix etc.) Methodology is adequately clear Hypothesis is adequately validated via the results and discussions.	Report is very well written and presented in a clear and concise manner. Contains well-presented amount of data exploration and explanation. (e.g. plot of histogram, correlation matrix etc.) Methodology is presented clearly and convincing Hypothesis are validated and correlates well with the results, and highly convincing.
Report Evaluation 2: (Technical Content, 25 marks) • **Discusses all processes involved to solve the problem.** • **Presents the background theory clearly for the reader for the different techniques used in solving the problem.** • **Experimental protocol is well defined and presented (usage of confusion matrix or other metrics to determine the performance)**	Technical content meets the bare minimal standard required for a technical report. Results are presented at a bare minimal. Source code of the demo is commented at a minimal level.	Technical content presents well on all the methodologies, processes and contains enough background knowledge to help user understand the overall solution to the problem. Results are well presented and accompanied with performance metrics (e.g. confusion matrix, area under the ROC curve etc.) Source code is well structured and commented.	Technical content is of highest quality and presents excellently on all the methodologies, processes and contains enough background knowledge to help user understand the overall solution to the problem. Results are well presented with well accompanied performance metrics and limitations of the results are well discussed (e.g. confusion matrix, area under the ROC curve etc.) Source code is of high quality and well

• **Commented source code**			commented as well as structured.

Deliverables:

- Report & PPTX & 5 mins video
- Deliverables detailing the solution to the problem that your group had chosen.

3   Hackathon Timeline

- Kick-off Week: Week 6 (Deadline: 14/4/2021)
  - o Register group members.
  - o Initial design of the problem containing problem statement and proposed hypothesis.
  - o Gantt Chart and tasks distribution between members
- Submission of Assignments: Week 12
- Voting of projects: Week 12 – Week 13
- Pitching Day: Week 14

Cash Prizes

1. First Place: RM350
2. Second Place: RM100
3. Third Place: RM 50

Note:

- Evaluation of all assignments for all groups, shall be done by me (i.e., single point of evaluation) based on Rubric 2 using the assignments submitted in Week 12.
- Winners of hackathon shall earn bragging rights.